IMDB MOVIES ANALYSIS

- ## Project Description

  The large dataset has given which contains the details of imdb movies like Director name, Movie name, budget, gross, actors, reviews and many more.

  Based on this dataset some analysis has done by using Microsoft Excel. The work is quiet complex and full of excitements.
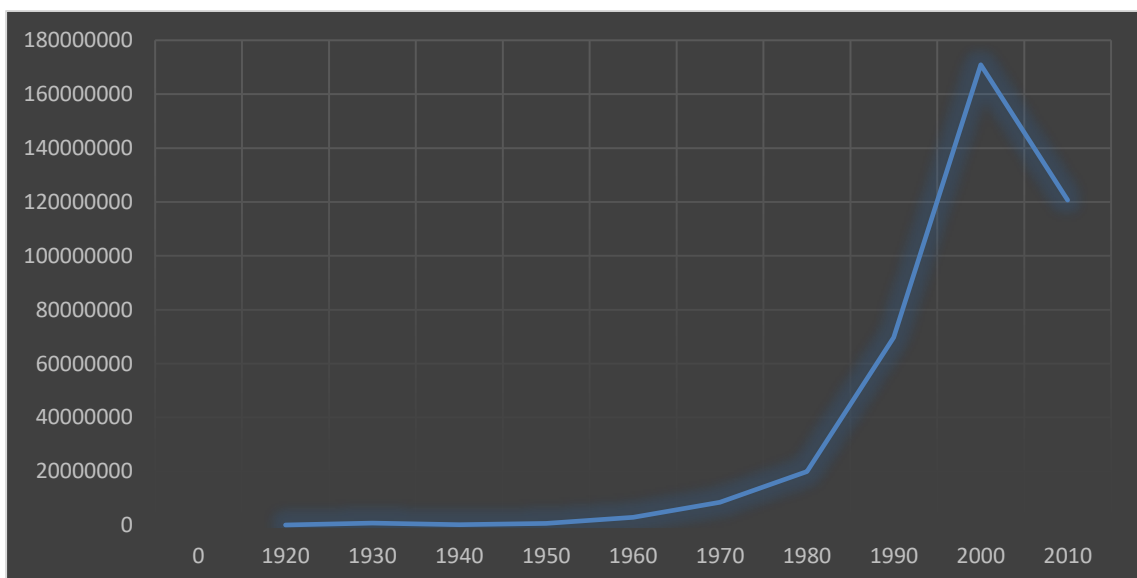
- ## Tech stack

  Microsoft Excel 2021

- ## Approach

  The first and basic step while analysing the data is to load the data into software which we are going to use. After this, explored the data and understand it according to the requirements. The next step is data cleaning and then go for the further analysis.

- ## Result

  The graph of decade vs total number of voted users calculated through excel:

- ## **Task 1**

  Clean the data

  In this step,

  1. Remove the Duplicates
  2. Remove the unnecessary columns
  3. Remove the blanks from the columns
  4. Replace the Null values with appropriate values in Language colum
  5. Fix the formatting of gross and budget columns.

- ## **Task 2**

  Find the top 10 movies with the highest profit?

  In this,

  1. Created new column "Profit" by using gross and budget columns
  2. Sort the Profit column in descending order and highlight the top 10 movies.

director_name ▾	num_critic_for_reviews ▾	gross ▾	actor_1_name ▾	movie_title
James Cameron	723	$ 76,05,05,847	CCH Pounder	Austin Powers: The Spy Who Shagged MeÂ
Colin Trevorrow	644	$ 65,21,77,271	Bryce Dallas Howard	The Twilight Saga: Breaking Dawn - Part 2Â
James Cameron	315	$ 65,86,72,302	Leonardo DiCaprio	The IncrediblesÂ
George Lucas	282	$ 46,09,35,665	Harrison Ford	TransformersÂ
Steven Spielberg	215	$ 43,49,49,459	Henry Thomas	TedÂ
Joss Whedon	703	$ 62,32,79,547	Chris Hemsworth	Crocodile DundeeÂ
Roger Allers	186	$ 42,27,83,777	Matthew Broderick	Pirates of the Caribbean: The Curse of the Black PearlÂ
George Lucas	320	$ 47,45,44,677	Natalie Portman	Pretty WomanÂ
Christopher Nolan	645	$ 53,33,16,061	Christian Bale	Guardians of the GalaxyÂ
Gary Ross	673	$ 40,79,99,255	Jennifer Lawrence	Dances with WolvesÂ

- ## **Task 3**

  Find the IMDB Top 250 Movies?

  ## **Result**

  1. Sort the movies which have more than 25000 voted users.
  2. After this sort same data in descending order by IMDB score.
  3. And then extract the top 250 movies from it and filter the movies which are in "English".
  4. Assign the rank to all the top 250 movies.

Top IMDB 250 Movies

Rank	director_name	num_critic_for_reviews	gross	genres	actor_1_name
1	Frank Darabont	199	$ 2,83,41,469	Crime|Drama	Morgan Freeman
2	Francis Ford Coppola	208	$ 13,48,21,952	Crime|Drama	Al Pacino
3	Christopher Nolan	645	$ 53,33,16,061	Action|Crime|Drama|Thriller	Christian Bale
4	Francis Ford Coppola	149	$ 5,73,00,000	Crime|Drama	Robert De Niro
5	Peter Jackson	328	$ 37,70,19,252	Action|Adventure|Drama|Fantasy	Orlando Bloom
6	Quentin Tarantino	215	$ 10,79,30,000	Crime|Drama	Bruce Willis
7	Steven Spielberg	174	$ 9,60,67,179	Biography|Drama|History	Liam Neeson
8	Sergio Leone	181	$ 61,00,000	Western	Clint Eastwood
9	Robert Zemeckis	149	$ 32,96,91,196	Comedy|Drama	Tom Hanks
10	Irvin Kershner	223	$ 29,01,58,751	Action|Adventure|Fantasy|Sci-Fi	Harrison Ford
11	Peter Jackson	297	$ 31,38,37,577	Action|Adventure|Drama|Fantasy	Christopher Lee
12	Christopher Nolan	642	$ 29,25,68,851	Action|Adventure|Sci-Fi|Thriller	Leonardo DiCaprio
13	David Fincher	315	$ 3,70,23,395	Drama	Brad Pitt
14	George Lucas	282	$ 46,09,35,665	Action|Adventure|Fantasy|Sci-Fi	Harrison Ford
15	Peter Jackson	294	$ 34,04,78,898	Action|Adventure|Drama|Fantasy	Christopher Lee
16	Lana Wachowski	313	$ 17,13,83,253	Action|Sci-Fi	Keanu Reeves
17	Milos Forman	149	$ 11,20,00,000	Drama	Scatman Crothers
18	Martin Scorsese	192	$ 4,68,36,394	Biography|Crime|Drama	Robert De Niro
19	Fernando Meirelles	214	$ 75,63,397	Crime|Drama	Alice Braga
20	Akira Kurosawa	153	$ 2,69,061	Action|Adventure|Drama	Takashi Shimura

- ## **Task 4**

  Find the top 10 best directors?

  ## **Results**

  In this,

  1. Created a new pivot table.
  2. Drop the director name column into Rows box of pivot
  3. Drop the Imdb score column into values box of pivot and calculated the average of imdb scores for directors.
  4. Sort the table on the basis of Average of Imdb Score.
  5. Highlight the top 10 directors.

Top 10 Directors

S.no	Director_name	Average of imdb_score
1	Charles Chaplin	8.6
2	Tony Kaye	8.6
3	Alfred Hitchcock	8.5
4	Damien Chazelle	8.5
5	Majid Majidi	8.5
6	Ron Fricke	8.5
7	Sergio Leone	8.4
8	Christopher Nolan	8.4
9	Asghar Farhadi	8.4
10	Marius A. Markevicius	8.4

- **Task 5**

Find the most popular genres?

Result

In this,

1. Delimited the Genres Column on the basis of "|" symbol by using Text to Column feature of Excel.
2. Created a new column "Popular genres" and store the first two genres.
3. Created a new pivot table.
4. Drop the popular genres column into rows box and gross column into values box of pivot and calculated the average of gross.
5. Sort the data on the basis of average of gross.

10 Most Popular Genre

Movie Genres	Average of gross
Family-Sci-Fi	$ 43,49,49,459
Adventure-Action	$ 31,75,57,891
Adventure-Sci-Fi	$ 19,51,95,016
Action-Documentary	$ 18,15,24,468
Musical-Comedy	$ 18,13,60,000
Animation-Horror	$ 14,83,13,048
Biography-Horror	$ 14,10,60,506
Drama-Comedy	$ 13,65,62,971
Horror-Comedy	$ 12,89,25,216
Animation-Drama	$ 12,76,02,381

- **Task 6**

Find the Critic favourite and Audience favourite actor?

Results

Leonardo di Caprio is the audience favourite actor.

In this,

1. Created a new pivot table.
2. Drop the Actor1 column into rows box and filter the data of "Brad Pitt", "Leonardo di Caprio" & "Meryl Streep"
3. Drop the "Critic Reviews" and "user reviews" column into values box and calculated the average of both of them
4. Sort the data on the basis of user reviews in descending order.

Row Labels ↓T	Average of num_critic_for_	Average of num_user_for_reviews
Leonardo DiCaprio	330.19	914.48
Brad Pitt	245.00	742.35
Meryl Streep	181.45	297.18
Grand Total	267.24	716.18

- **Decade vs Number of Voted Users (Bar graph)**